

Session - 11

SAGEMAKER

Oct 31, 2025



AGENDA

- What is Sagemaker
- How to train & deploy ML models

Different Ways of deploying ML models??

① Streamlit: When I want to showcase to client

② Flask + NGINX: For exposing API for my backend from so they can integrate it in application.

③ ECS → Production grade deployment
10 LAKH req.
3- 5- 4 min latency

④ EC2 → Production environment with
Low usage & steady access.

Location Type	Region				
AWS Region	Asia Pacific (Mumbai)				
Select an operating system, instance type, and vCPU to view rates					
Operating system					
Linux					
Instance type	vCPU				
All	All				
Viewing 604 available instances					
c5.4x					
Instance name	On-Demand hourly rate	vCPU	Memory	Storage	Network performance
c5.4xlarge	\$0.68	16	32 GiB	EBS Only	Up to 10 Gigabit

5

Sagemaker: For exposing API for my backend team so they can integrate it in application..

More Convenient.

Why do we need Sagemaker??

Recommendation Engine : download 1 level low.
Forbidden

download any code-base
Not allowed

Sagemaker : allows you to create jupyter notebooks

Inside AWS Infra

→ S3

→ ECR

→ Code-commit (GitHub alternative for AWS)

Google GCP → Virtren AI
 Azure → Azure ML
 Databricks → Databricks

A screenshot of the AWS SageMaker Canvas interface. The top navigation bar includes the AWS logo, a search bar containing "sageMaker Canvas", and various icons. The left sidebar has sections for "Amazon SageMaker" (Getting started, Dashboard, What's new), "Application Studio" (Studio, Canvas), and "Services" (Services, Features, Documentation, Knowledge articles, Marketplace, Blog posts, Events, Tutorials). A red arrow points from the "Services" menu item in the sidebar to the "Amazon SageMaker AI" service card in the main content area.

Services

- Amazon SageMaker**
The center for data, analytics, and AI
- Amazon SageMaker AI**
Build, Train, and Deploy Machine Learning Models
- Simple Queue Service**
SQS Managed Message Queues

A screenshot of the Amazon SageMaker AI landing page. The top navigation bar includes the AWS logo, a search bar containing "Search", and account information ("United States (Oregon) devops"). The left sidebar has sections for "Amazon SageMaker AI" (Getting started, Dashboard, What's new), "Applications and IDEs" (Studio, Canvas, RStudio, TensorBoard, Profiler, Notebooks, Partner AI Apps), and "Admin configurations" (Domains). A blue arrow points from the "Notebooks" link in the sidebar to the "Notebooks" section on the landing page. The main content area features a large heading "Amazon SageMaker AI" with the subtext "Build, train, and deploy machine learning models at scale". It also includes a subtext "The quickest and easiest way to get ML models from idea to production." and a "New to SageMaker AI?" section with "Set up for single user" and "Set up for organizations" buttons.

Amazon SageMaker AI

Build, train, and deploy machine learning models at scale

The quickest and easiest way to get ML models from idea to production.

New to SageMaker AI?

Quick setup for a single user ⓘ
This is perfect for first time users to try capabilities in just a few clicks.

Set up for single user

Advanced setup for organizations
Customize capabilities, permissions, network, and more for your team to launch Studio.

Set up for organizations

aws Services Search for services, features, blogs, docs, and more [Option+S] Oregon

Amazon SageMaker > Notebook instances > Create notebook instance

Create notebook instance

Amazon SageMaker provides pre-built fully managed notebook instances that run Jupyter notebooks. The notebook instances include example code for common model training and hosting exercises. [Learn more](#)

Notebook instance settings

Notebook instance name Maximum of 63 alphanumeric characters. Can include hyphens (-), but not spaces. Must be unique within your account in an AWS Region.

Notebook instance type

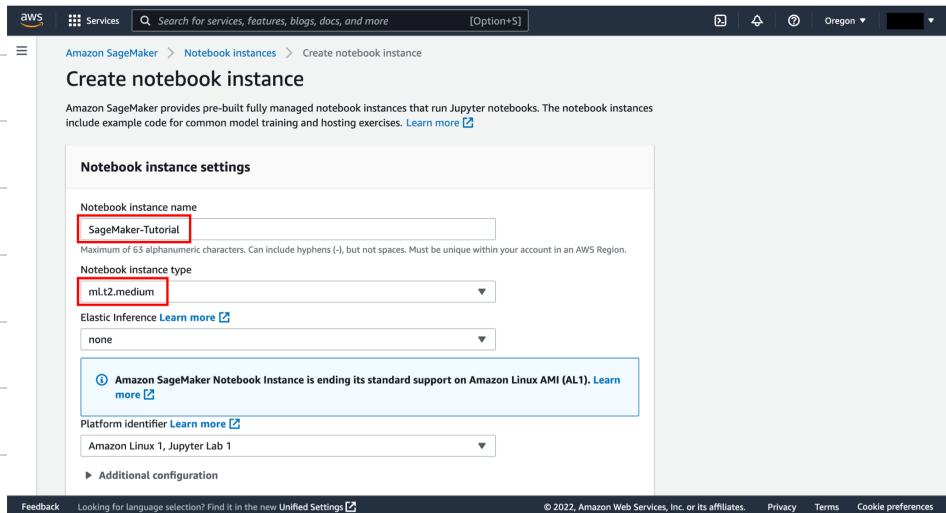
Elastic Inference [Learn more](#)

Platform identifier [Learn more](#)

① Amazon SageMaker Notebook Instance is ending its standard support on Amazon Linux AMI (AL1). [Learn more](#)

▶ Additional configuration

Feedback Looking for language selection? Find it in the new [Unified Settings](#). © 2022, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences



Jupyter

Files Running Clusters SageMaker Examples Conda

Select items to perform actions on them.

The notebook list is empty.

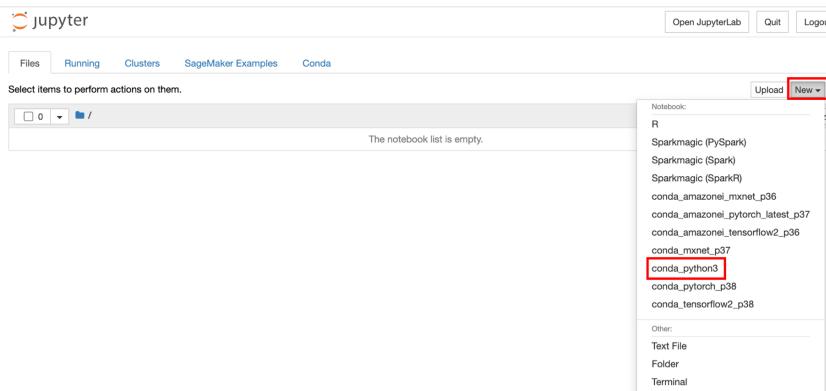
Upload [New](#)

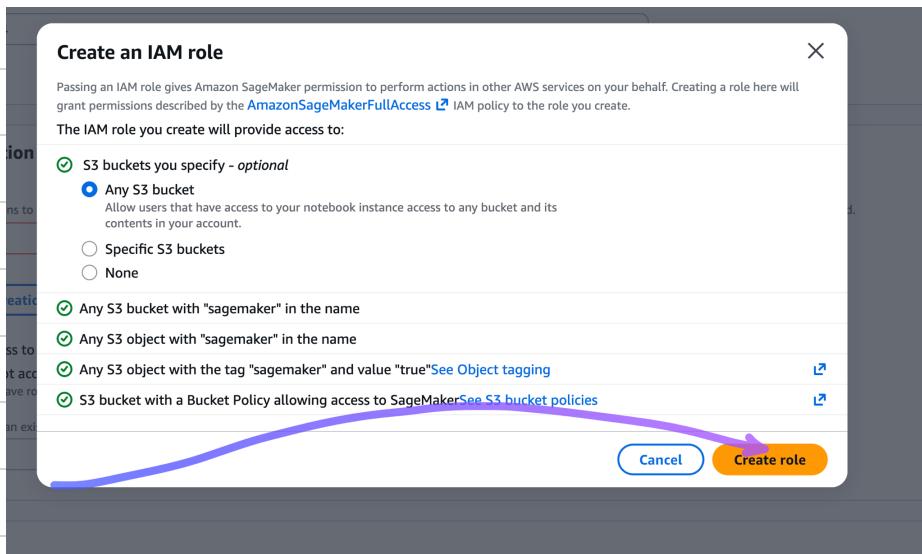
Notebook:

- R
- Sparkmagic (PySpark)
- Sparkmagic (Spark)
- Sparkmagic (SparkR)
- conda_amazonel_mxnet_p36
- conda_amazonel_pytorch_latest_p37
- conda_amazonel_tensorflow2_p36
- conda_mxnet_p37
- conda_python3**
- conda_pytorch_p38
- conda_tensorflow2_p38

Other:

- Text File
- Folder
- Terminal





Digressing to S3 now, for creating bucket

Storage

Amazon S3

Store and retrieve any amount of data from anywhere

Amazon S3 is an object storage service that offers industry-leading scalability, data availability, security, and performance.

Create a bucket

Every object in S3 is stored in a bucket. To upload files and folders to S3, you'll need to create a bucket where the objects will be stored.

[Create bucket](#)

How it works

Introduction to Amazon S3 | Amazon Web Services

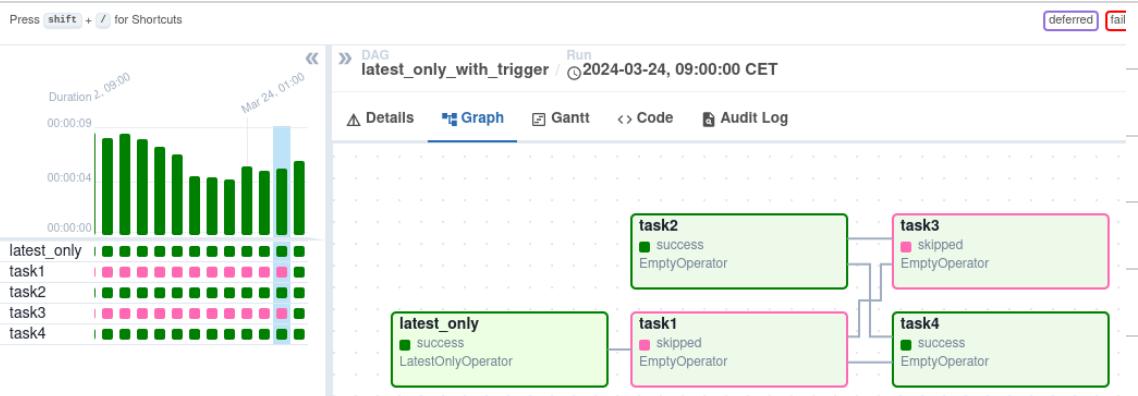
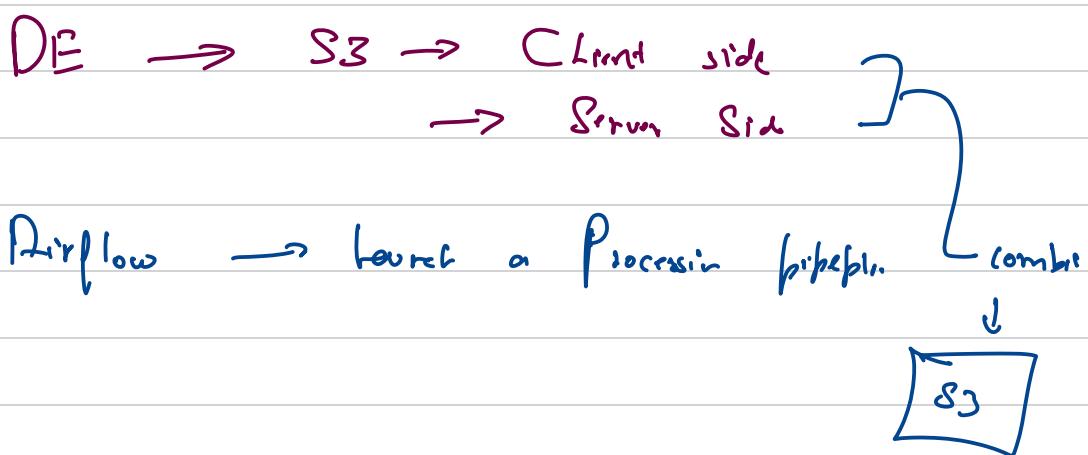
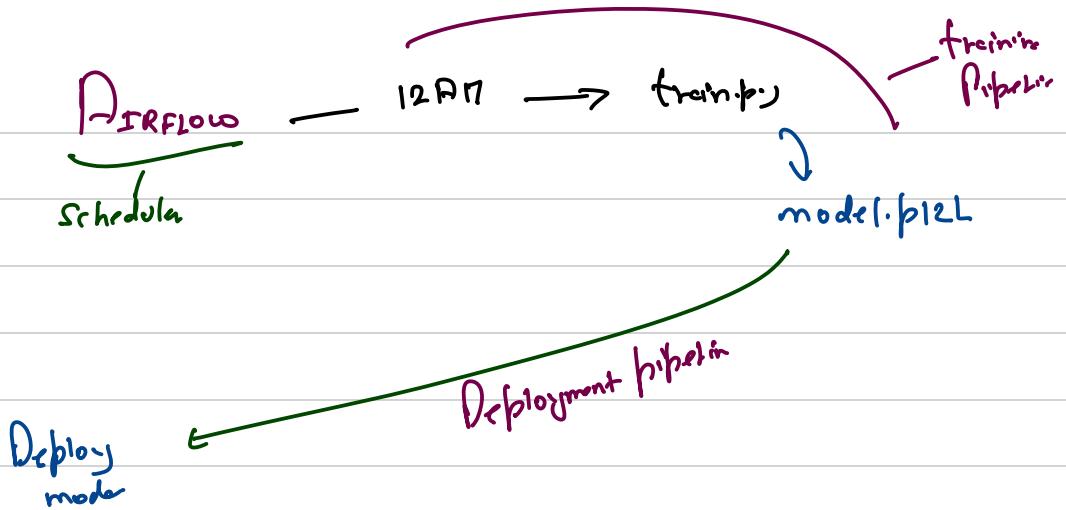
Copy link

Pricing

With S3, there are no minimum fees. You only pay for what you use. Prices are based on the location of your S3 bucket.

Estimate your monthly bill using the [AWS Simple Monthly Calculator](#)

[View pricing details](#)



What is Amazon SageMaker primarily used for?

4 options

Active Duration (Most preferred: 30 seconds)

Appears for 60 Secs

A To provide website hosting services

B As a chatbot interface for customer support

C To enable developers to create, train, and deploy ML models

D To facilitate large-scale data storage solutions

Data is not updated from the session



Aditya Pokhriyal
2/1 ⚡ 97.43



Rajesh Sharma
1/1 ⚡ 97.43



Anoop
3/1 ⚡ 95.70

4	SP	Surya Prakash	1/1 ⚡ 94.82
5	VY	Vinay Yelanadu	1/1 ⚡ 93.73
6	L	lakshmi santhi	1/1 ⚡ 92.96
7	AG	Aniket Gulhane	1/1 ⚡ 92.27
8	SA	Saurabh Arunkumar Gupta	1/1 ⚡ 89.97
9	RT	Riddhi Tatke	1/1 ⚡ 88.36
10	SC	sarita Chauhan	1/1 ⚡ 86.23

In Sagemaker

1

Create notebook ✓

2

Train a model.

3

Deploy a model.

4

Special function.

From hosting data

S3

Slow notebook container

Deploy

Running Instance — For EDA

Training Instances. — For training

①

P2. 16x longer — \$14.4

3 - hours on EDA/exploration — 14.4×3

0.5 → model training — 14.4×0.5
 $\approx \$50.4$

②

m5.2n long — EDA — \$0.4

3 - hours on EDA/exploration — $\$0.4 \times 3$

0.5 → model training — $\$14.4 \times 0.5$
 $\approx \$8.4$

How does Amazon SageMaker simplify the process of starting with machine learning models?

0 users have participated

A	By automatically writing code for the developer	0%
B	By letting developers choose the machine type without managing AMIs or security groups	0%
C	By requiring developers to manually manage machine instances and security	0%
D	By solely allowing the use of local machines for model training	0%

[End Quiz Now](#)



What allows Amazon SageMaker to balance ease of use and access to powerful computing resources for machine learning model development?

0 users have participated

A	Its ability to limit model training to only small datasets	0%
<input checked="" type="checkbox"/> B	The flexibility to select machine types and manage resources without complex configurations	0%
C	The option to use Amazon's proprietary programming language exclusively	0%
D	A fixed suite of tools that prescribe a specific workflow for all ML projects	0%

[End Quiz Now](#)

